Revue française d'informatique et de recherche opérationnelle. Série verte

R. G. Cassidy
C. A. Field
M. J. L. Kirby

Random payoff games with partial information : one person games against nature

Revue française d'informatique et de recherche opérationnelle. Série verte, tome 5, n° V3 (1971), p. 3-17
<http://www.numdam.org/item?id=RO_1971__5_3_3_0>

R.I.R.O.
(5e année, V-3, 1971. p. 3-18)

RANDOM PAYOFF GAMES WITH PARTIAL INFORMATION : ONE PERSON GAMES AGAINST NATURE

by R. G. CASSIDY [1], C. A. FIELD and M. J. L. KIRBY [2]

Abstract. — *In this paper we analyze the game situation in which the player has incomplete or partial information concerning the random payoff from the game. Several criteria for such a decision maker are formulated with solutions based on the techniques of linear programming. Our results extend considerably the previous results which had been obtained for the solution of random payoff games with partial information, and give the optimizing player a higher expected payoff from the game.*

I. INTRODUCTION

By a random payoff game with complete information we mean a two-person, zero sum game with mxn payoff matrix $A = \{ a_{ij} \}$, where A is a random variable with known distribution function $F(A)$. a_{ij} represents the payoff from player II to player I when player I plays row i and player II plays column j. Since a_{ij} is a random variable, the actual payoff on any play of the game will be $a_{ij}(w)$ where w is selected from the domain of a_{ij} according to the known marginal distribution of a_{ij}.

Since a_{ij} is a random variable a player does not known with certainty what the outcome of a particular play of the game will be even if he knows what pure strategy his opponent will use on that particular play. Thus, because of the randomness of the payoff matrix, both players are, in essence, forced to gamble. Under these circumstances it may be reasonable to use an optimality criterion other than the minimax criterion of deterministic zero sum games. For this reason several different concepts of optimality were introduced in [2] and [3].

One optimality criterion considered in [2] is for a player to maximize the probability, α, of his attaining a given, or prescribed, payoff level, β, no matter

(1) Now at Carnegie-Mellon University, Pittsburgh, Pennsylvania.
(2) Dalhousie University, Halifax, Nova Scotia, Canada.

what strategy his opponent plays. If this is players *I's* objective, then the problem of finding his optimal mixed strategy X can be formulated as the following linear programming problem.

$$(1.1)\qquad \begin{aligned} &\max \alpha \\ &\text{S.T.} \sum_{i=1}^{m} x_i P(a_{ij} \geqslant \beta) \geqslant \alpha \qquad \forall j \\ &\sum_{i=1}^{m} x_i = 1 \\ &x_i \geqslant 0 \qquad \forall i \end{aligned}$$

where $P(a_{ij} \geqslant \beta)$ is the probability that the random variable a_{ij} is greater than or equal to β. [2] contains a detailed analysis of model (1.1).

One generalization of (1.1) which we use below consists of having player I specify a range $[\beta_1, \beta_2]$ of desired payoff levels, β, and a weight function $w(\beta)$ which assigns an « a priori » weight to each payoff level β in $[\beta_1, \beta_2]$. Using this model the player determines his optimal mixed strategy X by solving the problem :

$$(1.2)\qquad \begin{aligned} &\max \alpha \\ &\text{S.T.} \sum_{i=1}^{m} x_i \int_{\beta_1}^{\beta_2} P(a_{ij} \geqslant t) w(t)\, \mathrm{d}t \geqslant \alpha \ \forall j \\ &\sum_{i=1}^{m} x_i = 1 \\ &x_i \geqslant 0 \qquad \forall i \end{aligned}$$

In [2] we have shown that if all a_{ij} have a finite range then β_1, β_2 can be chosen so that $w(t) = \dfrac{1}{\beta_2 - \beta_1}$ for all t in $[\beta_1, \beta_2]$ and (1.2) becomes the expected value game, that is, the deterministic-zero sum game in which each a_{ij} is replaced by its mean.

In both models (1.1) and (1.2), the fact that the distribution function $F(A)$ is known by the players means that the marginal distribution functions of each a_{ij} can be computed, and hence $P(a_{ij} \geqslant \beta)$ is known for any given value of β. However, in many problems the distribution function of the payoff matrix A may not be known and hence $P(a_{ij} \geqslant \beta)$ cannot be computed. In this chapter we consider the problem of solving random payoff games when the players have only partial information about the distribution function of the payoff matrix.

We begin by assuming that the payoff matrix A has only a finite number of possible states of nature. Thus we are assuming that each a_{ij} is a discrete random variable which can take on only a finite number of values. We let $A(k)$, $k = 1, 2, \ldots K$ be the k^{th} possible state of nature of A. Thus $A(k)$ is an mxn matrix. We define $P(k)$ by $P(k) = P(A = A(k))$.

Assumptions must now be made concerning the type of information the players have about the distribution function of A.

We consider the following degrees of information in a random payoff game :

a) The players have no information about the distribution function $F(A)$.

b) The players have a partial ordering on the probabilities of the states of nature of A.

c) The players know bounds on the probabilities of the various states of nature. That is, the players know that $a_k \leqslant p(k) \leqslant b_k$ where a_k, b_k are given, $k = 1, 2, \ldots K$.

d) The players have complete information about the distribution $F(A)$.

Cases (*a*) and (*d*) represent the extremes with respect to the amount of information available to the players. The standard method of solving (*a*) is to play a normal minimax game against nature (see [5] and [7] for example). This provides a conservative method of decision making when faced with total uncertainty. Case (*d*) can be solved in several different ways depending on the goals of the players. This has been the subject of study in [2] and [3]. Case (*b*) was introduced by Fourgeaud et al in [5]. There the case of a one person random payoff game against nature is analyzed and a method of solution for such a game is given. The results of [5] are generalized in section II below. Case (*c*) represents another type of partial information available to the players which we feel is very realistic in some settings. This case is also analyzed below. In addition we give an analysis and outline methods of solution for problems where the partial information is a combination of cases (*b*) and (*c*).

II. SUMMARY AND EXTENSION OF EXISTING RESULTS

Consider a one person random payoff game against nature with mxl payoff matrix $A = \{ a_i \}$. A is assumed to be a random matrix which has K possible states of nature, $A(k)$, $k = 1, 2, \ldots, K$. Let $a_i(k)$ be the i^{th} element of $A(k)$.

Following the development in [5], we assume that each player has some information about the probability that A takes on the value $A(k)$ and that this information takes the form of a partial ordering on the states of nature of A. The partial ordering can be expressed in the following way :

state $k \geqslant$ state k' (or $k \geqslant k'$) if the state of nature k has at least as great a probability of occuring as the state of nature k'.

Let P be any probability measure on the states of nature of A. Then P can be expressed as the vector $P = (p(1), \ldots, p(K))$ with $p(k) \geqslant 0$ and $\sum_{k=1}^{K} p(k) = 1$.

We say that the probability measure P is compatible with the given partial ordering if :

$$k \geqslant k' \Rightarrow p(k) \geqslant p(k')$$

for all partially ordered pairs k and k'. We define the set of compatible probability measures to be :

$$\Pi = \{ P \mid P \text{ is a probability measure and compatible with } \geqslant \}$$

Since the partial ordering results in a finite number of inequalities of the form $p(k) \geqslant p(k')$, Π is a convex polyhedron.

In [5] it is assumed that the player's optimality criterion is to maximize the expected value of his payoff. This means that the player would play the deterministic game with payoff matrix $\{ E(a_i) \}$ if he had enough information to calculate the expected values $E(a_i)$. The question then arises as to how the player should play the game when he does not have enough information to calculate $E(a_i)$. In [5] the authors assume that in this case the player will play as conservatively as possible. This latter assumption means that in place of $E(a_i)$, the player will use the minimum possible value of $E(a_i)$ over all probability measures $P \in \Pi$. Thus the player replaces $E(a_i)$ by M_i where

$$M_i = \min_{P \in \Pi} \sum_{k=1}^{K} p(k) a_i(k)$$

The player's optimal strategy is then that pure strategy i_o, for which M_{i_o} represents a lower bound for the player's expected payoff. By using the model developed below, the player will be able to improve this lower bound on his expected payoff.

To develop our model we use the optimum criterion of (1.2). Thus we suppose that the player would solve (1.2) if he had enough information to calculate the probabilities $P(a_{ij} \geqslant \beta)$ for all β in $[\beta_1, \beta_2)$.

That is, the player wants to maximize his probability of attaining a payoff in the interval $[\beta_1, \beta_2]$ where he assigns each payoff β in the interval the a priori weight $w(\beta)$. Since $n = 1$ we can ignore the second subscript which appears in (1.2). Thus the problem the player wants to solve can be written as :

$$\begin{aligned} & \max \alpha \\ & \text{S.T.} \sum_{i=1}^{m} x_i \left[\int_{\beta_1}^{\beta_2} P(a_i \geqslant t) w(t) \, \mathrm{d}t \right] \geqslant \alpha \\ (2.1) \qquad & \sum_{i=1}^{m} x_i = 1 \\ & x_i \geqslant 0 \qquad \forall i \end{aligned}$$

However, for any choice of $[\beta_1, \beta_2]$

$$\int_{\beta_1}^{\beta_2} P(a_i \geqslant t) w(t)\,\mathrm{d}t = \sum_{k=1}^{K} \left(\int_{\beta_1}^{\beta_2} \varphi_{ik}(t) w(t)\,\mathrm{d}t \right) p(k)$$
$$= E(\tilde{a}_i)$$

$$\text{where} \quad \varphi_{ik}(t) = \begin{cases} 1 & \text{if } a_i(k) \geqslant t \\ 0 & \text{otherwise} \end{cases}$$

and $\tilde{a}_i$ is a random variable which on the k^{th} state of nature, takes on the value $a_i(k) = \int_{\beta_1}^{\beta_2} \varphi_{ik}(t) w(t)\,\mathrm{d}t$ with probability $p(k)$. Thus (2.1) can be considered to be an expected value problem with $E(\tilde{a}_i)$ in place of $E(a_i)$.

Moreover since (2.1) has the same form as an expected value problem, the techniques developed in [5] can be carried over directly to (2.1). In particular if we assume, as in [5], that when the player does not have enough information to calculate $E(\tilde{a}_i)$ he will play as conservatively as possible, then in place of the unknown $E(\tilde{a}_i)$ in (2.1), the player will use the minimum possible value of $E(a_i)$ over all probability measures $P \in \Pi$. Thus in (2.1), $E(\tilde{a}_i)$ would be replaced by

$$\tilde{M}_i = \min_{P \in \Pi} \sum_{k=1}^{K} p(k) \tilde{a}_i(k)$$

As stated in the introduction, an appropriate choice of β_1, β_2 and the choice $W(t) = \dfrac{1}{\beta_2 - \beta_1}$ for all $t \in [\beta_1, \beta_2]$ gives $\tilde{M}_i = M_i$ so (2.1) is a direct generalization of the model considered in [5].

In order to summarize the results from [5] concerning methods of computing $M_i = \min_{P \in \Pi} \sum_{k=1}^{K} p(k) a_i(k)$, we must introduce the definition of an admissible support $S(P)$ of a probability measure P. We define

$$S(P) = \{ k \mid p(k) > 0 \}$$

and

$$S(\Pi) = \{ S(P) \mid P \in \Pi \}$$

$S(\Pi)$ is called the set of admissible supports for Π, and there are $2^k - 1$ of them. Then from [5, p. 15] we have the following theorem :

Theorem 2.1

$$M_i = \min_{P \in \Pi} \sum_{k=1}^{K} p(k) a_i(k)$$
$$= \min_{S(P) \in S(\Pi)} \frac{1}{|S(P)|} \sum_{k \in S(P)} a_i(k)$$

where $| S(P) |$ = cardinality of $S(P)$.

This theorem carries directly over to (2.1) and we have the result that

$$\tilde{M}_i = \min_{S(P)\in S(\Pi)} \frac{1}{|S(P)|} \sum_{k\in S(P)} \int_{\beta_1}^{\beta_2} \varphi_{ik}(t) w(t)\, dt$$

In [5] theorem (2.1) is used as the basis of a numerical procedure for solving (2.2) $\max_i \min_P \sum_{k=1}^{K} p(k)a_i(k)$ S.T. $P \in \Pi$. We now show that this theorem results from the fact that the optimal solution to the linear programming problem (2.2) must occur at one of the extreme points of the convex set of feasible solutions Π.

Theorem 2.2

If $S(\Pi)$ contains N admissible supports $S_1, S_2, \ldots, S_N$ where $N \leqslant 2^k - 1$ then the set of probability measures P_j

$$j = 1, 2, \ldots N \quad \text{where} \quad P_j(k) = \begin{cases} 1/|S_j| & \text{if } k \in S_j \\ 0 & \text{if } k \notin S_j \end{cases}$$

contains the extreme points of Π.

Proof : In order to prove this we show that any probability measure $P \in \Pi$ other than $P_j, j = 1, 2, \ldots, N$ is not an extreme point of Π.

Consider an arbitrary measure $P \in \Pi$, $P \neq P_j$, $j = 1, \ldots, N$ and such that P has $n \leqslant K$ non zero elements. By relabelling the states if necessary we can assume that $p(1) \geqslant p(2) \geqslant \ldots \geqslant p(n) > 0$. Since

$$P \neq P_j,\ j = 1, 2, \ldots, N, p(1) > p(n).$$

Let i_1 be the smallest integer such that

$$p(1) > p(i_1 + 1)$$

Choose

$$\epsilon = \min\,(1 - p(1), p(n), p(1) - p(i_1 + 1))$$

By the definition of i_1, $\varepsilon > 0$.

Let

$$Q = (p(1) - \left(\frac{n - i_1}{n}\right)\epsilon, \ldots p(i_1) - \left(\frac{n - i_1}{n}\right)\epsilon, p(i_1 + 1) + \frac{i_1}{n}\epsilon, \ldots p(n) + \frac{1}{n}\epsilon$$

We now show that $Q \in \Pi$. To see this, note that any relation of the partial ordering which holds between $p(i)$ and $p(j)$ with $i \leqslant i$, and $j \leqslant i$, or $i \geqslant i$, and $j \geqslant i$, will still hold for Q since all elements of Q in this range have been changed by the same amount.

If a relation of the partial ordering requires that $p(i) \geqslant p(j)$ with $i \leqslant i_1$, and $j > i_1$, then

$$p(i) - \frac{(n-i_1)}{n} \epsilon \geqslant p(j) + \frac{i_1}{n} \epsilon$$

since $p(i) - p(j) \geqslant \epsilon$. Also clearly $q(i) \geqslant 0$ and $\sum_{i=1}^{n} q(i) = 1$ so that $Q \in \Pi$.

Let

$$R = \left(p(1) + \frac{(n-i_1)}{n} \epsilon, \ldots p(i_1) + \frac{(n-i_1)}{n} \epsilon, p(i_1+1) - \frac{i_1}{n} \epsilon, \ldots p(n) - \frac{i_1 \epsilon}{n} \right)$$

Then if the partial ordering requires that $p(i) \geqslant p(j)$, this relation will hold for the corresponding components of R since we are adding a constant to the larger $p(i)'s$ and subtracting a constant from the smaller ones. Also, $\sum_{i=1}^{n} r_i = 1$ and $r_i \geqslant 0$, so that $R \in \Pi$.

But $P = 1/2\, Q + 1/2\, R$ and hence is not an extreme point of Π.

Since in calculating M_i, we are solving a linear program and since in theorem 2.2 we have shown that the probability measures $P_j, j = 1, \ldots, N$ contain the extreme points of the constraint set Π, theorem 2.1 follows directly as a corollary of theorem 2.2.

III. ALTERNATE SOLUTION PROCEDURES FOR GAMES WITH PARTIAL ORDERING

In order to develop new models for the solution of a one person random payoff game against nature in which the player knows a partial ordering on the states of nature of the payoff matrix, we introduce the following definitions and notation.

Let $S_1, S_2, \ldots, S_N$ be the set of admissible supports. Let P_j be defined as in theorem 2.2. From theorem (2.2), we know that the set of measures

$$\{ P_1, P_2, \ldots, P_N \}$$

contains the extreme points of Π.

Let us now restrict our attention to the extreme points of Π, say $\{ P_1, P_2, \ldots, P_T \}$, where $T \leqslant N$. We will call $\{ P_1, \ldots, P_T \}$ the extreme probability measures of Π. Define :

$$b_{ij} = \sum_{k=1}^{K} a_i(k) p_j(k) \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, T.$$

Then for every extreme probability measure, P_j, of the set Π we have a corresponding mxl vector expected payoffs :

$$B(j) = \{ b_{ij} \}$$

We now show how the mxT payoff matrix $B = \{ B(1), ..., B(T) \}$ can be used to develop a new method of solution for a random payoff game.

In [5] the player used the objective function :

$$\max_{i} \min_{P \in \Pi} \sum_{k=1}^{K} a_i(k)p(k)$$

Thus the player was only permitted to use a pure strategy solution. If, instead, we allow the player to use a mixed strategy solution then the player determines his optimum mixed strategy by solving :

$$(3.1) \qquad \begin{aligned} &\max_{X} \min_{P} \sum_{i=1}^{m} x_i \left(\sum_{k=1}^{K} a_i(k)p(k) \right) \\ &\text{S.T.} \quad P \in \Pi \\ &\sum_{i=1}^{m} x_i = 1 \\ &x_i \geqslant 0 \qquad \forall i \end{aligned}$$

Since $\{ P_j \}, j = 1, 2, ..., T$ are the extreme points of Π, any probability measure, $P \in \Pi$ can be expressed as a linear combination of $P_j, j = 1, 2, ..., T$. That is, there exist $\lambda_j, j = 1, 2, ..., T$, such that :

$$P = \sum_{j=1}^{T} \lambda_j P_j \quad \text{where} \quad \sum_{j=1}^{T} \lambda_j = 1, \lambda_j \geqslant 0 \qquad \forall j$$

By substituting this expression for P into (3.1) and interchanging the order of summations in the objective function, we obtain the problem :

$$(3.2) \qquad \begin{aligned} &\max_{X} \min_{\lambda} \sum_{1=1}^{m} \sum_{j=1}^{T} x_i \left(\sum_{k=1}^{K} a_i(k)p_j(k) \right) \lambda_j = \sum_{i=1}^{m} \sum_{j=1}^{T} x_i b_{ij} \lambda_j \\ &\text{S.T.} \quad \sum_{i=1}^{m} x_i = 1 \\ &\sum_{j=1}^{T} \lambda_j \geqslant 0 \\ &x_i \geqslant 0 \qquad \forall i \\ &\lambda_j \geqslant 0 \qquad \forall j \end{aligned}$$

But (3.2) is simply a deterministic two person zero sum game with mxT payoff matrix

$$P = \{ b_{ij} \}$$

where X is the mixed strategy for player I and λ is the mixed strategy for player II. Moreover, since a choice of λ by player II yields a probability measure $P \in \Pi$ via the relation $P = \sum_{j=1}^{T} \lambda_j P_j$, we can think of player II in (3.2) as being nature, since in the development of the previous section $P \in \Pi$ is selected by nature. Thus (3.1) is equivalent to the player playing a deterministic two person zero sum game with payoff matrix B with nature as his opponent.

To solve (3.2) we rewrite it (see [6]) as the linear programming problem :

$$
\begin{aligned}
&\max \alpha \\
&\text{S.T.} \sum_{i=1}^{m} x_i b_{ij} \geqslant \alpha \qquad \forall j \\
&\sum_{i=1}^{m} x_i = 1 \\
&x_i \geqslant 0 \qquad \forall i
\end{aligned}
\tag{3.3}
$$

In solving the game whose payoff matrix is B, the player selects the *mixed* strategy X which maximizes his expected payoff when his opponent, nature, is allowed to choose a mixed strategy $P \in \Pi$ which is the most unfavourable strategy from the point of view of the player. This is in contrast to the method of solution given in [5] in which the player was only allowed to select a pure strategy. This leads to the following conclusions :

Theorem 3.1

Let

$$v_1 = \max_{i} \min_{\lambda} \sum_{j=1}^{T} b_{ij}\lambda_j$$

and

$$v_2 = \max_{X} \min_{\lambda} \sum_{i=1}^{m} \sum_{j=1}^{T} x_i b_{ij} \lambda_j$$

Then $v_1 \leqslant v_2$ and v_2 gives a lower bound on the expected payoff which the player will receive when he plays the random payoff game with partial information.

The importance of this theorem is that it shows precisely why we believe that the method of solution given in [5] is too pessimistic. It is assumed in [5] that the player would choose i_o to maximize $\{ E(a_i) \}$ if he had enough information to compute $E(a_i)$ $i = 1, 2, \dots m$. If the player has insufficient information to compute $E(a_i)$ and if he wishes to play conservatively then he should adopt a minimax optimality criterion in which he chooses his strategy X so as to maximize the worst possible expected payoff that the could get for any $P \in \Pi$. Thus the problem he should solve is (3.1) or equivalenty (3.2). Moreover, since $v_1 \leqslant v_2$, it follows that v_1, the optimal value of the objective function obtained from (2.2) gives a pessimistic or too conservative an estimate of the actual expected value the player will receive.

This can be seen more clearly in two other ways. First note that v_1 is the optimal value of (3.1) when the player is allowed to choose only a pure strategy solution rather than a mixed strategy solution. Thus v_1 will inevitably give a worse estimate of the actual expected payoff than v_2, since v_1 is obtained by solving a more constrained problem.

Alternatively note how v_1 was obtained. For each $i = 1, ..., m$ the player assumed nature would chose the $P \in \Pi$ which gave the minimum value of $\sum_{k=1}^{K} a_i(k)p(k)$. This means the player assumes, in effect, that nature will change its strategy whenever the player plays a different pure strategy.

But a probability measure $P \in \Pi$ applies to the whole vector $A(k)$ not just to the i^{th} row of it. Thus, in a realistic model, the player should assume nature chooses a strategy $P \in \Pi$ which will hold for all rows of A, not just the i^{th} row. But the model in [5] acts as if nature could change $P \in \Pi$ whenever the player's mixed strategy required that he play a different row from the one he played in the previous play of the game. This appears to be unduly pessimistic and, in fact, not in keeping with the basic assumption of the model, which is that there exists a true distribution $P \in \Pi$ which holds for the whole vector A.

To illustrate the results of Theorem (3.1) more clearly we present an example.

EXAMPLE 3.1

Consider a 4×1 game with four possible states of nature and the following partial ordering on the states of nature :

$$p(1) \geqslant p(2) \geqslant p(4), \quad p(3) \geqslant p(4)$$

The admissible supports are :

$$\{1\}, \{3\}, \{1, 3\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}$$

The extreme probability measures are $P_1 = (1, 0, 0, 0)$, $P_2 = (0, 0, 1, 0)$, $P_3 = (.5, .5, 0, 0)$, and $P_4 = (.25, .25, .25, .25)$.

Suppose that the values of the states of nature $a_i(k)$ are as given in the following table :

STATES $k =$	1	2	3	4
TRUE PROBABILITY	5/12	4/12	2/12	1/12
—	—	—	—	—
$a_1(k)$	3	3	3	.68
$a_2(k)$	3	1	1	4.32
$a_3(k)$	2	4	1	5
$a_4(k)$	2	4	3	3

The true or actual expected values for each state are : $E(a_1) = 2.98$, $E(a_2) = 2.11$, $E(a_3) = 2.73$, and $E(a_4) = 2.91$.

Corresponding to the four extreme probability measures P_1, P_2, P_3 and P_4 we have the following four vectors of expected payoffs :

$$B(1) = \begin{bmatrix} 3 \\ 3 \\ 2 \\ 2 \end{bmatrix} \qquad B(2) = \begin{bmatrix} 4 \\ 1 \\ 1 \\ 3 \end{bmatrix} \qquad B(3) = \begin{bmatrix} 3 \\ 2 \\ 3 \\ 3 \end{bmatrix} \qquad B(4) = \begin{bmatrix} 2.67 \\ 2.33 \\ 3 \\ 3 \end{bmatrix}$$

We then form the matrix $B = (B(1), B(2), B(3), B(4))$ and solve the two person zero sum game with payoff matrix B.

The optimal solution is $x_1 = .75, x_2 = 0, x_3 = 0, x_4 = .25$ and $\alpha = 2.75 = v_2$.However, if we use the solution method of [5] we find that $v_1 = 2.67$ and the optimal strategy is $x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0$. Thus the method of solution proposed in this section does indeed give a better lower bound on the actual expected payoff than the method of solution given in [1].

If we use the true probabilities to calculate the expected values $E(a_i)$, we find that by playing the pure strategy $X = (1, 0, 0, 0)$ the player would have an expected return of 2.98. Thus if the player had sufficient information to be able to calculate $\{ E(a_i) \}$ he would play row one all the time.

In section II we showed that the expected value optimality criterion was a special case of the optimality criterion in which the player tries to maximize the probability of achieving a payoff in the interval $[\beta_1, \beta_2]$ when each payoff level β in the interval is assigned an a priori weight $w(\beta)$. Moreover we also showed that the model with this generalized optimality criterion could be solved as an expected value model if we replaced a_i by the more general random variable $\tilde{a}_i$. We can now make this same generalization of the model considered in this section and consider instead of b_{ij} the quantity C_{ij} where

$$C_{ij} = \sum_{k=1}^{K} \left(\int_{\beta_1}^{\beta_2} \varphi_{ik}(t) w(t) \mathrm{d}t \right) p_j(k) \qquad \begin{matrix} j = 1, 2, \dots, T \\ i = 1, 2, \dots, m \end{matrix}$$

Thus instead of a payoff matrix B we can consider a deterministic two person zero sum game with payoff matrix $C = (C(1), C(2), ..., C(T)) = \{ C_{ij} \}$.

To illustrate this model we solve example (3.1) with $\beta = 3$, that is $w(t) = 1$ if $t = 3$ and 0 otherwise.

EXAMPLE 3.2

Using the data of the previous example and $\beta = 3$ we calculate C to be

$$C(1) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad C(2) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad C(3) = \begin{bmatrix} 1 \\ .5 \\ .5 \\ .5 \end{bmatrix} \quad C(4) = \begin{bmatrix} .75 \\ .5 \\ .5 \\ .75 \end{bmatrix}$$

To see how these values were obtained note that $C_{13} \geqslant P_3(a_1 \geqslant 3)$ where $P_3 = (.5, .5, 0, 0)$ and hence $C_{13} = .5 + .5 = 1$. In a similar manner we calculate the remaining entries.

We then solve the two person zero sum game with payoff matrix C. The optimal solution is $x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0$ and $\alpha = .75$. This optimal strategy will result in the player obtaining a payoff of 3 with a probability of 75 %.

It is important to note that, in the above development, the matrices B and C could be enlarged to include columns $B(i)$ and $C(j)$ for each admissible support $S(P_j)\, j = 1, 2, ..., N$ and their respective probability measures. That is, augmenting B and C by additional columns obtained by using all the probability measures $P_1, P_2, ..., P_N$ rather than just the extreme probability measures $P_1, P_2, ..., P_T$ does not in any way change the above results. This is because the essential property we have used, namely that any probability measure $P \in \Pi$ can be expressed as a convex combination of the extreme points of Π, continues to hold as long as the set of points considered contains all the extreme points as a subset. This is a well known game theory result that a payoff matrix can be increased by adding columns which are linear combinations of existing columns without changing the optimal strategy or value of the game. The only rationale for augmenting B and C in this way is that it avoids the computational problem of selecting the extreme probability measures, $P_1, P_2, ..., P_T$, from the set $P_1, P_2, ..., P_N$.

We now consider the effect of additional information in the solution of a random payoff game.

IV. FURTHER INFORMATION ON STATE PROBABILITIES

Suppose the players knowledge of the random matrix A consists of a partial ordering on the states of nature (as in the previous section) and of bounds on the probability of state k occuring. These bounds are assumed to be given in the form :

$$f_k \leqslant P(\text{state } k \text{ occuring}) = p(k) \leqslant g_k$$

In [5] and in sections II and III above the only case considered is the one in which $f_k = 0$ and $g_k = 1$ for all k. Note however that if $f_k = g_k$ for all $k = 1, 2, \ldots, K$, then we are in the situation of having complete information about the distribution of the random matrix A. The solution of this problem was discussed in [2] and [3].

Therefore the addition of the information on bounds on $p(k)$ enables us to consider the cases ranging from no information ($f_k = 0$, $g_k = 1$ all k) up to complete information ($f_k = g_k$, $k = 1, 2, \ldots, K$) about the distribution of A.

We assume that the bounds and the partial ordering information are compatible. That is, if $p(k) \geqslant p(k')$ then $g_k \geqslant f_{k'}$. Let Π be defined as before and let

$$\Phi = \{ P \mid f_k \leqslant p(k) \leqslant g_k, \forall\, k \}$$

We adopt the same optimality criterion as in section III. Thus the player assumes that nature will choose the most unfavourable distribution from its feasible set of probability distributions $P \in \Pi \cap \Phi$ and then the player maximizes his expected payoff against this probability distribution. Thus the player uses a maximum optimality criterion. This means that the player seeks a strategy X' such that :

$$\min_{P \in \Pi \cap \Phi} \sum_{i=1}^{m} \sum_{k=1}^{K} x_i' a_i(k) p(k) = \max_{X} \min_{P \in \Pi \cap \Phi} \sum_{i=1}^{m} \sum_{k=1}^{K} x_i a_i(k) p(k)$$

Thus he wants to find the optimal solution to :

$$\begin{aligned} & \max_{X} \min_{P} \sum_{i=1}^{m} \sum_{k=1}^{K} x_i a_i(k) p(k) \\ (4.1) \qquad & \text{S.T. } P \in \Pi \cap \Phi \\ & \sum_{i=1}^{m} x_i = 1 \\ & x_i \geqslant 0 \qquad \forall i \end{aligned}$$

This is a problem of exactly the same form as (3.1) and so could be solved by the same techniques of it were possible to find the extreme points of the convex polyhedron $\Pi \cap \Phi$. Unfortunately no simple way of finding the extreme points of $\Pi \cap \Phi$, or of finding a set of points which contains the extreme points appears to be available except in the special case discussed in section III where $\Pi \cap \Phi = \Pi$. Thus an alternate approach is needed.

One such approach is to regard (4.1) as a constrained deterministic two person zero sum game. In this game the two players have strategies X and P respectively, with player II, nature, having additional constraints on his set of possible mixed strategies. These additional constraints are in the form of linear inequalities since $P \in \Pi \cap \Phi$ can be written as a set of linear inequalities. This same constrained game approach could have been used for the solution of (3.1) in section III but we did not do so since it is easier to solve (3.1) by first finding the extreme probability measures of Π and then solving (3.3).

A method of solution for constrained games is given in [7] (§ 3.7) and [4]. We let the set of linear inequalities $P \in \Pi \cap \Phi$ be denoted by $PE \geqslant F$ where E and F are a matrix and vector of constants respectively, with the vector F containing the values f_k, g_k, $k = 1, \ldots, K$. Then the linear programming problem which is the equivalent to (4.1) is ([7]) :

$$\begin{aligned} &\max ZF^T \\ &\text{S.T. } ZE - XD \leqslant 0 \\ &Z, X \geqslant 0 \end{aligned} \tag{4.2}$$

Where $D = \{ a_i(k) \}$, and F^T is the transpose of F.

An interesting feature of (4.2) is that it indicates how a procedure for evaluating the value of perfect information might be developed by using a parametric programming analysis of the vector F and seeing how the optimal value of the objective function changes as f_k and g_k get close together.

It is interesting to note also that because the optimal strategy is found by solving a linear programming problem, linear constraints may be added to (4.2) without substantially increasing the difficulty of obtaining a solution. The following constraint which is linear in the $p(k)$ and involves the expected payoff against pure strategy i might be useful :

$$\sum_{k=1}^{K} p(k) a_i(k) \begin{smallmatrix} \geqslant \\ = \\ \leqslant \end{smallmatrix} \mu_i \tag{4.3}$$

where μ_i is the known mean for the distribution of a_i. In actual problems both the mean and variance of nature's distribution may be known, but to involve the following variance constraint :

$$\sum_{k=1}^{K} p(k) a_i^2(k) - \left(\sum_{k=1}^{K} p(k) a_i(k) \right)^2 = \sigma_i^2 \tag{4.4}$$

would make the problem non-linear in the $p(k)$'s and increase the computational difficulty substantially.

BIBLIOGRAPHY

[1] A. BAÑOS, « On Pseudo Games », *Ann. of Math. Stat.*, 39 (1968), 1932-1945.

[2] R. G. CASSIDY, « Random Payoff Games ». Ph. D. thesis, Dalhousie, University, 1970.

[3] R. G. CASSIDY, C. A. FIELD and M. J. L. KIRBY, « Random Payoff Games : A Payoff Maximization Model », submitted to *Management Science*, dec. 1970.

[4] A. CHARNES, « Constrained Games and Linear Programming », Proc. Nat. Acad. Sci. U. S. A. 39 (1953), 639-641.

[5] M.M. C. FOURGEAUD, B. LENCLUD and Ph. SENTIS, « Critere de Choix en Avenir Partiellement Incertain », *R.I.R.O.*, 3 (1968), 9-20.

[6] G. HADLEY, « Linear Programming », Addison Wesley, Reading, Mass., 1962.

[7] G. Owen, « Game Theory », W. B. Saunders, Philadelphia, 1968.